Genome-wide association studies in purebred and crossbred entire male pigs

Hélène GILBERT
helene.gilbert@toulouse.inra.fr

Objective

Genome wide association studies in purebreds and crossbreds
Are they common regions?
What could explain the differences?
Population structure and genotyping

Crossbred
Piétrain x LW mat
Piétrain x LW pat

Quality control
Illumina SNP chip
62,626 markers
42,379 purebred
47,247 crossbred
41,450 common markers

Purebreds
Piétrain
Traits

- **Production**: feed intake and growth (DFI, FCR, RFI, ADG), BW

- **Carcass composition**: Carcass yield, lean meat content, tomograph, ...

- **Meat quality**: a*, b*, L*, drip loss, ultimate pH

- **Lesions**: at mixing, before slaughter, on the carcass

- **Boar taint/sex hormones**: androstenone, skatole, indole, testosterone, oestradiol

- **Health parameters**: leucocytes, CRP, Pig-MAP, ...

- **Lameness**

78 traits
**Genome wide analyses**

**Linear mixed model**

\[ Y = \mu + X\beta + W\alpha + Zu + \varepsilon \]

- **Phenotypes**
- **Mean**
- **Fixed effects**
  - **Dosage matrix**
    - \(0 = 11\)
    - \(1 = 12\)
    - \(2 = 22\)
  - **SNP effect**
- **Polygenic effects**
  - \(u_j \sim N(0, G\sigma_r^2)\)
  - \(G = \text{genomic kinship}\)
- **Residual**

*Applied to purebreds and crossbreds separately*
GWAS (2) – log10(p.value) > 5, traits

1Mb regions – purebred results, 24 significant regions
1Mb regions – crossbred results, 18 significant regions

Lesions, posterior part, crossbreds
GWAS (3) common regions, trait families

For all SNP with $-\log_{10}(p\text{.value})>4$

In the 1 Mb-region surrounding the marker, look for $-\log_{10}(p\text{.value})>3$ in the trait family
GWAS (3) common regions, trait families

For all SNP with -log10(p.value)>4
In the 1 Mb-region surrounding the marker, look for -log10(p.value)>3 in the trait family

- 1Mb regions – purebred results, 325 significant regions
- 1Mb regions – crossbred results, 321 significant regions

3 to 17 regions per trait
GWAS (3) common regions, trait families

1Mb regions – purebred results, 325 significant regions

1Mb regions – crossbred results, 321 significant regions

Only 15 common regions:

- 6
- 5
- 3
- 1

Traits:
- Lameness
- Intake/Growth
- Meat color
- Drip loss
- Fat
- Sex hormones
- Lesions
- Muscle
- Boar taint
- Bones
- pHu
- BW
- Health parameters
GWAS - summary

1. Similar number of regions detected in purebreds and crossbreds

2. Only few common regions in the two populations
   a. Different allelic frequencies between breeds
   b. Different linkage disequilibrium (opposite marker/QTL phases)
   c. Different genomic background effect
   d. Limited power
Allelic breed effects

Breed origin of the alleles ➔ Are allelic effects different depending on the breed on crossbred animals?
Allelic breed effects

Breed origin of the alleles  Are allelic effects different depending on the breed on crossbred animals?

Large White

Q  q

1  2

Piétrain

Q  q

2  1

Dosage matrix

0 = 11
1 = 12
2 = 22

No effect
Allelic breed effects

Breed origin of the alleles → Are allelic effects different depending on the breed on crossbred animals?

Large White

$\begin{align*}
Q & \quad q \\
1 & = 11 \\
2 & = 12 \\
& = 22
\end{align*}$

Piétrain

$\begin{align*}
Q & \quad q \\
1 & = 11 \\
2 & = 12 \\
& = 21
\end{align*}$

Dosage matrix

No effect

Factor matrix

Is effect $2 \neq$ effect $3$?
Allelic breed effects

Are allelic effects different depending on the breed on crossbred animals?

- Large White
  - Dosage matrix:
    - 0 = 11
    - 1 = 12
    - 2 = 22
  - Factor matrix:
    - 1 = 11
    - 2 = 12
    - 3 = 21
    - 4 = 22
  - No effect

- Piétrain
  - Dosage matrix:
    - 2 = 11
    - 1 = 12
    - 2 = 22
  - Factor matrix:
    - 1 = 11
    - 2 = 12
    - 3 = 21
    - 4 = 22
  - Is effect 2 ≠ effect 3?

LW
- Is a_{LW} ≠ 0?

PI
- Is a_{PI} ≠ 0?
Allelic breed effects, 1Mb regions

Factor matrix
1 = 11
2 = 12
3 = 21
4 = 22

Factor matrix
930 regions

PI specific effect
38 regions

LW specific effect
43 regions

157 regions

5

5

Breed specific dosage matrices
\[ \begin{align*}
\text{LW} & \left[ \begin{array}{c}
0 = 1 \\
1 = 2
\end{array} \right] \\
\text{PI} & \left[ \begin{array}{c}
0 = 1 \\
1 = 2
\end{array} \right]
\end{align*} \]

6 to 29 regions per trait

1 to 24 regions per trait

6 to 29 regions per trait

1 to 24 regions per trait

Intake/Growth
Meat color
Fat
Sex hormones
Lesions
Muscle
Boar taint
Bones
pHu
BW
Health parameters
Drip loss
Lameness
Allelic breed effects, 1Mb regions

Factor matrix
1 = 11
2 = 12
3 = 21
4 = 22

Breed specific dosage matrices
\[
\begin{align*}
\text{LW} & \begin{bmatrix}
0 & 1 \\
1 & 2
\end{bmatrix} \\
\text{PI} & \begin{bmatrix}
0 & 1 \\
1 & 2
\end{bmatrix}
\end{align*}
\]

Factor matrix
930 regions

PI specific effect
157 regions

LW specific effect
43 regions

157 + 43 = 190 regions

6 to 29 regions per trait

1 to 24 regions per trait

Factor matrix
6 to 29 regions per trait

PI specific effect
157 regions

LW specific effect
43 regions

Breed specific dosage matrices

Intake/Growth
Meat color
Fat
Sex hormones
Lesions
Muscle
Boar taint
Bones
pHu
BW
Health parameters
Drip loss
Lameness
Allelic breed effects, 1Mb regions

Factor matrix
1 = 11
2 = 12
3 = 21
4 = 22

Breed specific dosage matrices
LW
\[ 0 = 1 \]
\[ 1 = 2 \]

PI
\[ 0 = 1 \]
\[ 1 = 2 \]

30% with different MAF in LW and PI

(MAF in LW) – (MAF in PI) > 0.2

42% with different MAF in LW and PI

Factor matrix
- 930 regions

PI specific effect
- 38 regions
- 157 regions

LW specific effect
- 43 regions
- 178 regions
Common regions across models

- purebreds and/or crossbreds
- factor matrix
- breed specific dosage matrices

Total of 2154 1Mb-regions detected with one model

291 were detected in at least two analyses
Common regions across models

- purebreds and/or crossbreds
  - factor matrix
  - breed specific dosage matrices

Total of 2154 1Mb-regions detected with one model

291 were detected in at least two analyses

0 to 7 common regions per trait

* Genetic correlations different from 1 between crossbreds and purebreds
Conclusions

- Not many regions detected in common between purebreds and crossbreds: about 40% of the traits had genetic correlations between purebreds and crossbreds different from 1

- Quite a lot of breed specific effects:
  - ~1/3 might be due to different allelic frequencies between breeds
    - What about the 2/3 left?
  - Show unequal distribution across trait families
    - Could relate to genetic correlations magnitudes between purebreds and crossbreds
Acknowledgements

From the test farm for records and samplings:
   A. Berre, L. Dantec, M.-H. Lohat, G. Martin, L. Udin

From Ifip for contributing to the design monitoring:
   C. Hassenfratz, A. Varenne

From the PEGASE research unit for laboratory analyses:
   R. Comte, S. Jaguelin, M. Lefebvre, F. Thomas

From the GenPhySE research unit for DNA preparation:
   K. Feve

This work is part of the national program UtOpIGe, carried out by INRA, IFIP, SYSAAF, Novogen, and Bioporc pig breeding companies (ADN, Choice Genetics France, Gène+, Nucleus)

Financial support was from ANR (ANR-10-GENOM_BTV-015 UtOpIGe), FranceAgrimer, Bioporc, InaPorc, and INRA (Integrated Management of Animal Health metaprogramme for the “GISA-HealthyGrowth” project)

M. Chassier is funded by the EU FP6 Integrated Project SABRE